Received: 19 January 2018 | Revised: 22 August 2018 | Accepted: 29 August 2018

DOI: 10.1002/gdj3.56

DATA SERVICES PAPER

WILEY Geoscience Data Journal RMetS

A roadmap to climate data rescue services

Stefan Brönnimann[1,2] | **Yuri Brugnara**[1,2] | **Rob J. Allan**[3] | **Manola Brunet**[4,5] | **Gilbert P. Compo**[6,7] | **Richard I. Crouthamel**[8] | **Philip D. Jones**[5,9] | **Sylvie Jourdain**[10] | **Jürg Luterbacher**[11,12] | **Peter Siegmund**[13] | **Maria Antónia Valente**[14] | **Clive W. Wilkinson**[5]

[1]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland
[2]Institute of Geography, University of Bern, Bern, Switzerland
[3]Met Office Hadley Centre, Exeter, UK
[4]Centre for Climate Change, Department of Geography, Universitat Rovira i Virgili, Tarragona, Spain
[5]Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, UK
[6]CIRES, University of Colorado, Boulder, Colorado
[7]Physical Sciences Division, NOAA Earth System Research Laboratory, Boulder, Colorado
[8]International Environmental Data Rescue Organization (IEDRO), Deale, Maryland
[9]Center of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz University, Jeddah, Saudi Arabia
[10]Météo-France, Direction de la Climatologie et des Services Climatiques, Toulouse, France
[11]Department of Geography, Climatology, Climate Dynamics and Climate Change, Justus Liebig University of Giessen, Giessen, Germany
[12]Center of International Development and Environmental Research, Justus Liebig University of Giessen, Giessen, Germany
[13]Royal Netherlands Meteorological Institute, De Bilt, Netherlands
[14]Instituto Dom Luiz, Faculdade de Ciências da Universidade de Lisboa, Portugal

Correspondence
Yuri Brugnara, Institute of Geography, University of Bern, Hallerstrasse 12, 3012 Bern, Switzerland.
Email: yuri.brugnara@giub.unibe.ch

Funding information
European Commission via Copernicus Climate Change Service (C3S), Physical Sciences Division of the NOAA Earth System Research Laboratory, Swiss National Science Foundation.

Abstract

Quantitative approaches to climate risk management such as mapping or impact modelling rely on past meteorological data with daily or sub-daily resolution, a large fraction of which have not yet been digitized. Over the last decade or so, a number of projects have contributed to the rescue of some of these data. Here we provide a summary of a survey we have undertaken of several meteorological and climate data rescue projects, in order to identify the needs of climate data rescue services. To make these efforts more sustainable, additional integrated activities are needed. We argue that meteorological and climate data rescue must be seen as a continuous, coordinated long-term effort. Technical developments (e.g. data assimilation), new scientific questions (e.g. process understanding of extreme events) and new social (e.g. risk assessment, health) or economic (e.g. new renewable energy sources, agriculture and forestry, tourism, infrastructure, etc.) services are highlighting the immense value of data previously neglected or never considered. This continuous effort is currently undertaken by projects of various sizes, structure, funding and staffing, as well as by dedicated programmes, ranging from those within many national weather services down to "grassroots" initiatives. These activities are often not sufficiently coordinated, staffed, or funded at an international level and will benefit considerably from climate data rescue services being established within the Copernicus Climate Change Service (C3S) (https://climate.copernicus.eu/).

KEYWORDS
C3S, climate change, climate data rescue, climate service, digitization

1 | INTRODUCTION

The past few decades have seen considerable progress in our scientific understanding and assessment of climate change and related consequences (IPCC, 2013). Adaptation and climate risk management are rapidly developing fields that increasingly use numerical models and rely on quantitative high resolution, high-quality and continuous data sources (e.g. Bowyer, *et al*., 2014; Challinor *et al*., 2018). With this comes the need for more and longer data series than that are currently available. In fact, the amount of information that could be made available, but exists in only hard copy form, is substantial. As a consequence of the growing need for data to support climate change research, climate data rescue has seen a marked revival over the past decade (Kwok, 2017). This resurgence has been a result of climate data rescue being pursued as a community effort (Allan *et al*., 2011a; 2011b), whose success depends not only on the individual groups, projects and initiatives involved, but on maintaining and growing a coherent and strong community to lead it forward.

Sustaining climate data rescue activities in this context might benefit from some organised form of support, which we term climate data rescue services. In this paper, through a systematic survey involving literature screening, a questionnaire on climate data rescue activities, plus our experience in various European and national projects, we have analysed what form of support best fits the needs of future climate data rescue activities. We provide a roadmap for meteorological and climate data rescue services, building on the roadmap for the storage and management of the meteorological data over land (Thorne *et al*., 2017) and the oceans (the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) version 3, Freeman *et al*., 2017).

In this paper, we first analyse how meteorological and climate data rescue currently works (an annex to this paper, which will be updated annually, provides an inventory of global activities) and from that derive three premises. From these follow requirements for climate data rescue services, embedded within a coherent, "end-to-end" IT infrastructure encompassing an international data portal, data registry, tools and techniques that are now being established by the Copernicus Climate Change Service within the 311a Lot1 for Collection and Processing of In Situ Observations - Data Rescue (abbreviated C3S DRS in the rest of the paper).

2 | CLIMATE DATA RESCUE

2.1 | Non-digitized data

Before sketching the current status of climate data rescue in Sect. 2.2, we first turn to the object of data rescue: the non-digitized data. In the context of this study, these are data that pre-date the age of electronic data acquisition. Especially in developing countries, weather observations are still performed manually and require digitization, which is not covered here. Neither are data in electronic format on obsolete media, which typically require different sort of services.

Although very long and important terrestrial surface and upper air data series have been digitized and processed (e.g. Alcoforado *et al*., 2012 for Portugal; Camuffo and Bertolin, 2012, Camuffo *et al*., 2013 and Camuffo *et al*., 2017 for Italy and the Western Mediterranean; Domínguez-Castro *et al*., 2014 for Spain; Brunet *et al*., 2014 and Ashcroft *et al*., 2018 for the Mediterranean North Africa and the Middle East; Slonosky, 2014 for Canada; Ashcroft *et al*., 2014 for southeastern Australia; Stickler *et al*., 2014 for global upper-air data, to give only very few examples), a huge fraction have still not been digitized (see also Bosilovich *et al*., 2013). This also applies to historical marine observations from exploration, naval, postal, merchant and passenger ships. Overall, the fraction of yet-to-be-digitized data is difficult to quantify. A wide range of repositories, such as the Ottoman, Austro-Hungarian and Venetian archives, are as yet little known to climate scientists (e.g. Luterbacher *et al*., 2012). Poorly-tapped data sources also include information contained in strip charts, e.g. thermographs and barographs, (Fig. 1), and a vast bulk of digitized meteorological and climate data from marine or terrestrial platforms, that has yet to be keyed from hard copy readings in one tabulated form or another, such as newspapers and published monographs.

Sources of historical weather observations from ship log books have almost exclusively focused on the holdings from the major European colonial powers (United

Kingdom, The Netherlands, Spain, Portugal, Germany—and are far from exhausted) and the United States. However, recent discoveries in various archives and museums in Norway and Finland have recovered some 53 K images of log books, with at least another 10 K log books from the latter country yet to be imaged (Wilkinson and Vásquez, 2016; Wilkinson and Vásquez, 2017). Archives in countries such as Sweden and Denmark have yet to be "explored" and those in France and Portugal are in need of greater attention (see Jourdain *et al.*, 2015). It is very important to note that many of these logs cover exploration and commercial/mercantile voyages into the "data sparse" oceans of the South Hemisphere middle to higher latitudes (see Figure 1 in Freeman *et al.*, 2017). The above examples only begin to illustrate the huge amount of non-digitized climate information that can be found in various archives around the world.

Most developed countries already have the infrastructure (organization of the archives, formation of the personnel, etc.) to efficiently carry out a national climate data rescue programme, although they rarely allocate enough resources for a full-time implementation. In developing countries, the World Meteorological Organization (WMO) and other organizations have been funding numerous programmes in recent years to reorganize archives, buy digitization hardware and software, and train the personnel, avoiding the loss of many documents to improper storage (e.g. Posada *et al.*, 2018). Many examples can be found on the WMO supervised International Data Rescue (I-DARE) Portal (https://idare-portal.org), where actual and potential climate data rescue activities can be reported (Siegmund, 2014). WMO also publishes guidelines for best practices in the digitization of climate data (WMO, 2016).

2.2 | Current status of climate data rescue: A survey

Who is doing the climate data rescue work presently and why? One of the tasks of C3S DRS is to answer these questions, in order to improve the coordination between different projects and avoid repetition of work. An online survey was launched in late 2017, targeted at the coordinators of climate data rescue projects. This exercise will be repeated every year to compile an evolving inventory of climate data rescue activities that is as complete as possible. Other than a general description of the activity, we asked in particular what is the geographical domain (which region and surface), which parameters are rescued and at what is their temporal resolution. The 30 respondents to the survey were representative of all six WMO regions, although the majority (60%) were from Europe.

Additional sources of information for the inventory were international meetings, including workshops of the Atmospheric Circulation Reconstructions over the Earth (ACRE; www.met-acre.org/) initiative, meetings of the European Geosciences Union, the European Meteorological Society and others, the scientific literature and the internet. In addition, we draw from our experience in a large number of previous projects and initiatives, in which the authors were or are involved (European projects European and North Atlantic daily to Multidecadal climate variability (EMULATE), Historical Instrumental Climatological Surface Time Series Of The Greater Alpine Region (HISTALP), Uncertainties in Ensembles of Regional Re-Analysis (UERRA), European Reanalysis And Observations for Monitoring (EURO4M), European Reanalysis of Global Climate Observations (ERA-CLIM1, ERA-CLIM2), community initiatives ACRE and International Surface

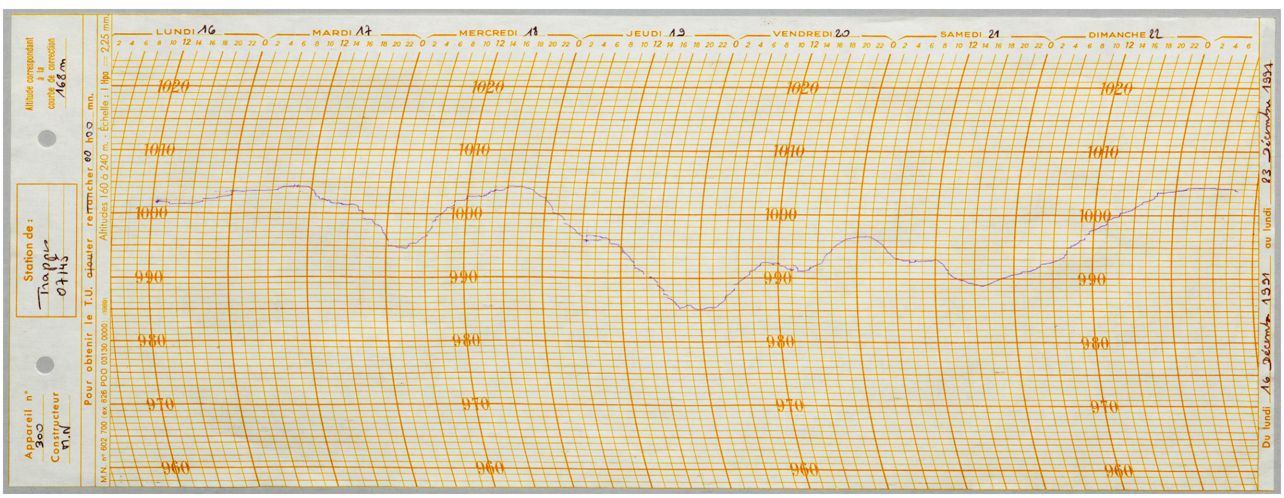

FIGURE 1 Barograph strip chart from a French station. Strip charts are a precious but often neglected source of high-resolution observations. Image credit: Météo-France

Temperature Initiative (ISTI) Climate Science for Service Partnership (CSSP) China, University of Giessen digitization initiative, Swiss projects Digitization and Homogenization (DIGIHOM) and Swiss Early Instrumental Measurements for Studying Decadal Climate Variability (CHIMES) and many more).

The inventory is given in the Supporting Information S1 and will be used to update the WMO I-DARE and the C3S DRS portals.

2.2.1 | Who rescues?

The survey results and the inventory show that climate data rescue is usually approached on three different time scales: (a) efforts of short duration (weeks or months), where the data to be rescued are determined by the specific needs of one-off activities (such as a research article, an undergraduate or graduate thesis) and carried out by a single person or research group; (b) projects with limited duration, usually 1-3 years, that are funded by regional, national or international institutions (such as the European Commission) to achieve predetermined goals that are usually tightly linked with research activities involving different institutions; (c) long-term endeavours, typically carried out by public national/regional weather services or through voluntary work (e.g. citizen science) and international 'grassroots' data community initiatives (e.g. ACRE). Our inventory gives three recent examples of these types: (i) the digitization of the London sub-daily pressure series from 1,692, coordinated by Richard Cornes (KNMI/CRU); (ii) the COBECORE project (https://cobecore.org), a 4-year project funded by the Belgian Science Policy Office with the aim to recover eco-climatological data in the Congo basin (Meeus, 2018); (iii) the climate data rescue activities performed regularly in the European national weather services, which are summarised on the EuMetNet—DaRe website (https://www.zamg.ac.at/dare/), where many weather services provide inventories of digital available data as well as non-digitized data existing in their archives.

Projects and one-off activities (time-scale type i and ii) can set priorities that better suit the needs of climate research and leave clear traces of what has been done by publishing reports or journal articles; moreover, in most cases the rescued data are immediately made public without restrictions. However, their short durations mean the loss of specific know-how that will take time to be acquired by a new project. An additional concern with such activities is the long-term curation of the rescued data (see Hsu *et al*., 2015). Long-term endeavours (type iii) allow a more efficient and thorough data rescue, but they can be hampered by a lack of funds or personnel. Moreover, in the case of national or regional weather services, the rescued data may be subject to limitations that embargo or curtail their distribution. In fact, there exists a certain amount of data that have already been digitized years or even decades ago, but which remains inaccessible to most of the scientific community. Other issues are obsolete digital data formats and poor or unsuccessful transfers between old and new data mediums.

In some cases, data owners may not even be aware that their data can be useful to others, or they do not know any platform where they can share them. It is common for rescued data or metadata to be stored on private computers until they are eventually forgotten or lost (particularly for one-off activities). This not only applies to keyed data, but also to digital images containing data and paper records. For example, during the WMO DARE1 programme in the 1990s, nearly 2 million pages of data from 48 African countries were imaged and placed on nearly 100,000 microfiche sheets. These sheets, originally held by the Belgian National Meteorological Service, were turned over to the African Centre of Meteorological Application for Development (ACMAD; located in Niamey, Niger) in the mid-2000s to be held in safe keeping until funding was available to digitize these imaged data. Unfortunately, although placing imagery onto microfiche and microfilm was a state-of-the-art technology at the time, scientists soon realized that exposed to temperatures close to 40°C for more than a decade, the images would soon fade. Compounding the problem is that many of the African countries destroyed the original paper records, as they were told that the information was now safe. Although a digitization project for the microfiche was eventually funded by the US Agency for International Development in 2013, this allowed for the scanning of only a small fraction of the sheets. At this point in time, neither ACMAD nor the WMO is any closer to rescuing and digitizing these historic surface data than when the DARE1 programme started. This example shows how climate data rescue can be negatively affected by the postponement of funding, as media such as paper and film degrade over time.

Nevertheless, several climate data rescue activities have delivered, and continue to deliver, data to the scientific community. A new focus for the climate data rescue community has seen the linking together of historical climatologists, historians and geographers with impact researchers (e.g. Dupigny-Giroux and Mock, 2009, Allan *et al*., 2011a; 2011b; 2016; Williamson et al., 2015; Klein et al., 2018; Williamson et al., 2018). At the centre of this activity is the previously mentioned international ACRE initiative (Figure 2). ACRE undertakes and facilitates historical global surface terrestrial, upper-air and marine weather data recovery, imaging and digitization. It provides these data to open-access international repositories responsible for such material, seeing that they provide the best quality and quantity of surface weather observations for assimilation into all reanalyses.

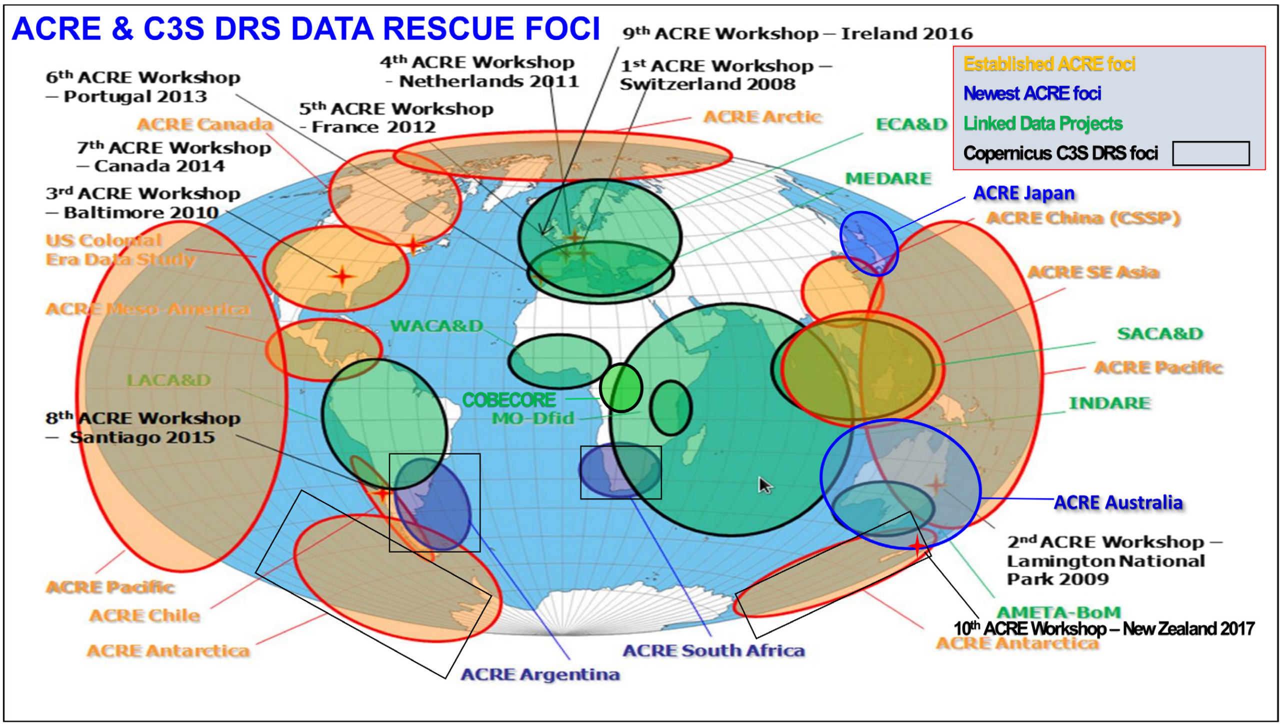

FIGURE 2 The interplay between the international ACRE initiative (including the locations of its various annual international workshops), linked data rescue projects and C3S DRS regional data rescue activities (updated from Allan *et al*., 2011a)

Furthermore, ACRE works to ensure that reanalysis outputs are freely available and feed seamlessly into the climate science, climate applications and services, impacts, risks and extremes communities (Allan *et al*., 2011a).

2.2.2 | What is rescued?

Most climate data rescue activities focus on data from terrestrial surface stations (Figure 3, top). They provide long (multi-decadal) records for the same geographical location and have, therefore, tremendous value in climate change research. However, marine data rescue is also vital, especially to the improvement of global reanalyses, as marine platforms can provide critical observations for remote regions or where land data are insufficient (tropics, southern hemisphere, polar regions), particularly in early colonial times. Reanalyses also benefit from upper-air data (Stickler *et al*., 2014), which is also instrumental in understanding climate change (e.g. Seidel *et al*., 2011; Thorne *et al*., 2011) and from sub-daily observations (Ashcroft *et al*., 2018). In the polar regions, observations involving ice surfaces and extents are of great importance (e.g. Turner and Comiso, 2017). Temperature, precipitation and pressure are the three core variables in climate data rescue (Figure 3, middle). While temperature and precipitation are often directly analysed, pressure is usually assimilated into atmospheric reanalyses (see summary in Fujiwara *et al*., 2017) or used for other types of large-scale circulation reconstruction (e.g. Ansell *et al*., 2006; Küttel *et al*., 2010; Schwander *et al*., 2017). Other variables, previously considered "minor," are now also being rescued regularly. Wind and radiation, in particular, have increasingly gained attention over the last decade in response to the observed "global stilling" (Vautard *et al*., 2010) and "global brightening" (Wild, 2012) phenomena. These variables are also particularly important to the emerging renewable energy sector (see e.g. C3S European Climate Energy Mixes and CLIM4ENERGY, https://ecem.climate.copernicus.eu, https://clim4energy.climate.copernicus.eu).

Awareness of the importance of individual observations is now widespread in the scientific community, therefore climate data rescue is no longer limited to daily or monthly means (Figure 3, bottom), as it was once (e.g. Jones *et al*., 1999; Camuffo and Jones, 2002). One-third of the respondents to the survey were dealing with hourly or finer resolutions (e.g. strip charts). The C3S DRS survey also included an open question on what assistance is needed. Nearly all of the respondents asked for resources for data keying. In fact, many projects can only afford the imaging of documents and leave the transcription of the data to future initiatives. With the involvement of volunteers in science made more and more attractive by modern technology (Bonney *et al*., 2014), it is easy to predict that citizen science will play an increasingly important role for climate

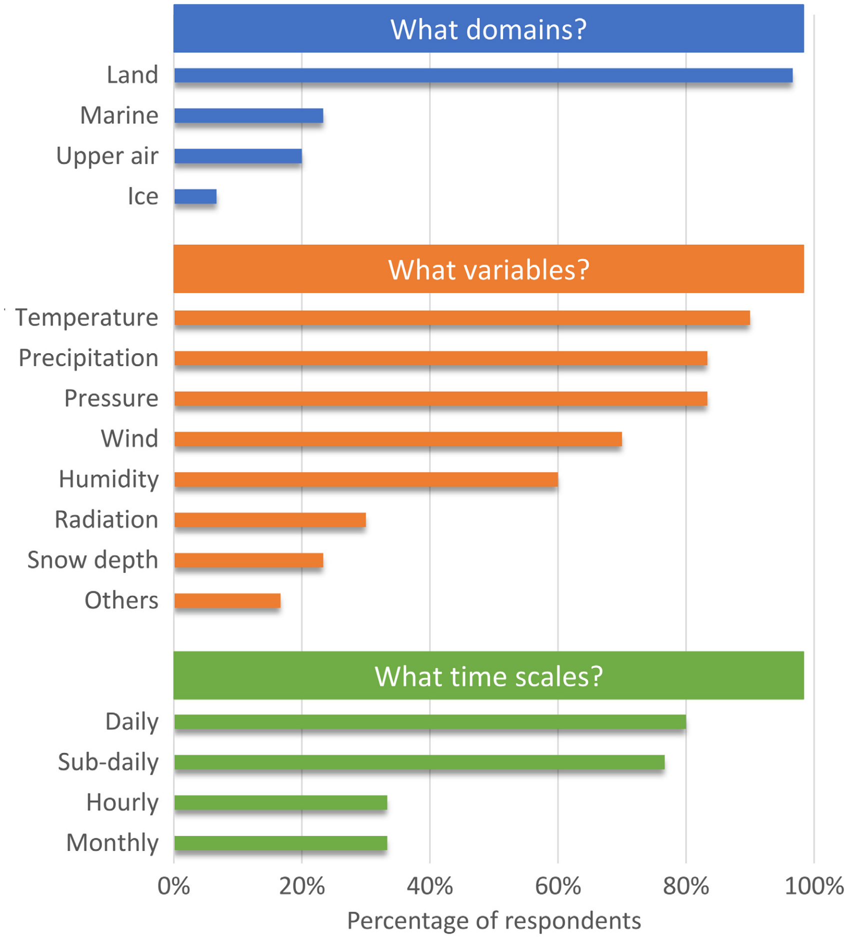

FIGURE 3 Survey results (based on 30 responses)

data rescue (10% of the respondents already rely mainly on citizen science for data keying).

3 | PREMISES FOR ADVANCING CLIMATE DATA RESCUE

Based on these results, we formulate three premises, which we then target in climate data rescue services.

3.1 | Climate data rescue as a continuous effort

The survey shows that temperature, precipitation and pressure are the main variables now being rescued, but increasingly other variables are also targeted. Additionally, it shows that sub-daily data are now often digitized whereas previously daily or monthly means were considered sufficient. The increased demand for historical weather data is at least partly due to new techniques, which have permitted the reconstruction of past global- to regional-scale weather patterns through the assimilation of such observations in ways that were not previously possible. Data assimilation approaches allow for the use of sparse observations to obtain useful new data products (Compo *et al.*, 2006; Compo *et al.*, 2011). As a result, several historical dynamical weather reconstructions or reanalyses spanning a century or more have been produced in the last few years (Compo *et al.*, 2011; Poli *et al.*, 2016; Laloyaux *et al.*, 2018), demonstrating that large-scale dynamical weather reconstruction is generally possible and can be improved by more observations from additional data rescue efforts. These long historical reanalyses primarily rely on sub-daily pressure observations, which previously have received less attention or interest by both climate scientists and historical climatologists. The demand is also driven by the changing focus of climate science from mean conditions towards extreme events. This confluence shows that new techniques (data assimilation), new scientific questions (understanding extreme events) and new social (risk assessment) or economic issues (new renewable energy sources) render data valuable that had hitherto not been considered so. Such changes in demand are likely to continue. Perhaps future long reanalysis efforts will make use of wind data over land, cloud motion or cloudiness (Toride *et al.*, 2017), rainfall, or soil temperature data. Likewise, some variables may contain important information on other properties of interest, such as aerosol properties that can possibly be derived from sunshine duration records (Sanchez-Romero *et al.*, 2014). The numerous spectroscopic data from solar and stellar spectra that were obtained in the early 20th century might eventually also be used for atmospheric science (e.g. Griffin, 2006), but only once suitable methods have been developed.

These new, important demands require going back to original archives and data repositories for the recovery, imaging, and digitizing of the relevant historical records. They also require evolving quality control and reprocessing to meet the need for high-quality sub-daily data. Thus, we need to consider weather and climate data rescue as a continuous long-term effort that evolves with new techniques and new demands, and we need to develop an international environment that can sustain this continuity (including the provision of proper infrastructures and catalogues and the retention or reference to the original records as much as possible).

3.2 | A distributed approach to climate data rescue

Although it is a global, continuous effort, it has been argued that climate data rescue is best performed by many small, individual projects (see Brunet and Jones, 2011) or by a mix of regional foci under a broad international umbrella (Allan *et al.*, 2011a; 2011b; 2016). Each of these projects or foci establish their own links to archives (e.g. Veale *et al.*, 2017), libraries, scientific and end users and stakeholders. This is confirmed by our survey, which shows that about 80% of the projects are small, targeting mostly one country or even just one station. National projects (such as Kaspar *et al.*, 2015) may be long-lasting and target national data holdings. International projects perform climate data rescue on a wider scale, but remain limited to 3-4 year efforts unless, as in the international ACRE initiative, they can engender ongoing commitment from the international climate data rescue community. Large masses of data can be rescued in citizen science projects, especially if they are well defined, such as oldweather.org (Figure 4). This is suitable for some of the sources that are well structured and whose context is well known. However, many other sources require dedicated scientific work and are not suitable for such efforts. The steps preceding the digitization (archive search, imaging) as well as the following steps (controlling, documenting, formatting, etc.) are typically more time-consuming and require expertise that is beyond most the scope of citizen scientists. Climate data rescue comprises a much wider range of tasks than scanning millions of pages and typing numbers from sheets.

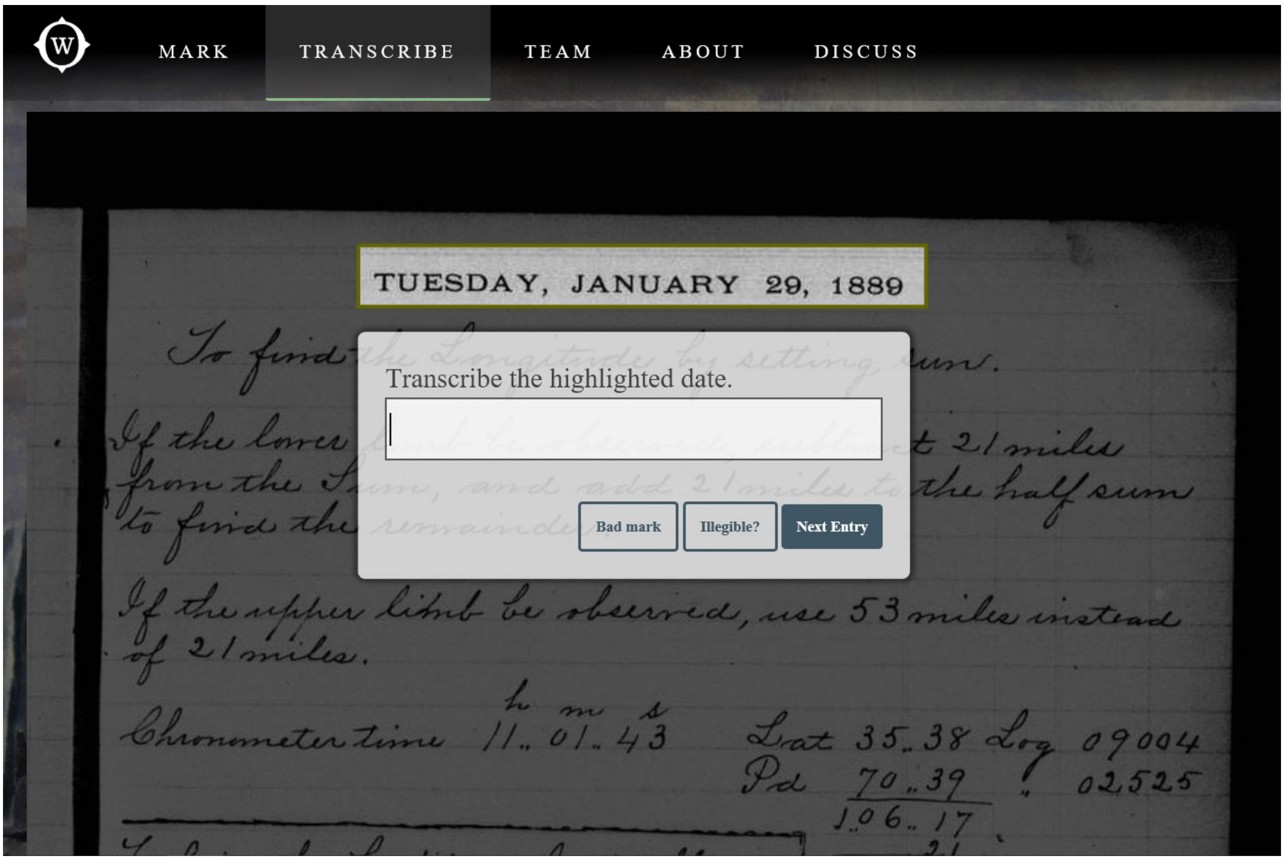

FIGURE 4 Screenshot of oldweather.org (new Whaling Chapter). Citizen science projects such as oldweather.org have provided millions of valuable ship data, but cannot “do the job”. Only together with research projects, efforts of national weather services and projects supported by development agencies and foundations can the body of climate data be rescued

The distributed approach also makes sense because local context is important. Not only is climate data rescue dealing with geographical data, which often requires knowledge of the past location and place names, but it is also dealing with historical data, which often requires knowledge and expertise about the organization of a local archive. Collecting this context, in turn, is important in itself as a source of information on climate-society interaction (Brönnimann and Wintzer, 2018).

At the same time, there needs to be close collaboration between climate data rescue groups, with networks shared where possible, including know-how on citizen science approaches, imaged library holdings and databases in humanities. What the ACRE initiative (Allan *et al*., 2011a; 2011b; 2016) has demonstrated is that an international umbrella can provide overarching linkages between various projects, programmes, institutions and organizations operating on various spatial scales—local, regional and global. It thus builds a coherent community, effectively "pooling" resources, personnel and funds from the various regional climate data rescue foci under it (e.g. Brunet *et al*., 2014; Kaspar *et al*., 2015; Williamson *et al*., 2015; Ashcroft *et al*., 2016; Posada *et al*., 2018, Williamson *et al*., 2018), and has been shown to be a viable role model for ongoing data rescue activities. The ACRE example suggests that climate data rescue requires maintaining an active, vibrant community.

3.3 | A multi-level effort

Our third premise is that efforts must target projects, the scientific community and science administration. For a distributed approach to work, it must be beneficial not only to society, but also to scientists. Or, as Griffin (2015: 97) puts it:"Data-rescue projects differ in a fundamental way from most research projects, where prestige and success are strong drivers. By contrast, data rescue is a domain for collaboration, not competition and in the broadest of senses." For an individual research project to engage in climate data rescue, there must be scientific merit that is acknowledged by the scientific community. Funding bodies should realise the importance of this work for science and the public, thus increasing the rate of acceptance for funding. Focusing on simplifying the required effort, such as through best practice guides, metadatabases and software tools, can reduce the needed effort, personnel, and resources. In turn, from the point of view of the scientific community, it is important that the individual projects are guided towards delivering data in a common format to the various responsible data repositories. The C3S DRS serves the needs both of projects and of the community, as is outlined in Section 11.

At the level of science policy and science administration, effective science requires an open, traceable data policy and the value of climate data science should be acknowledged. During the past 10 to 20 years, much has been achieved in working towards these aims. Datasets can obtain digital object identifiers and, in general, funding bodies are more willing to fund climate data rescue projects, especially when they generate new, high-quality databases and products that aid weather/climate applications and services communities. Several data journals have appeared that cater to the needs of the climate data rescue (and other) communities. Examples include the "Geoscience Data Journal," "Earth System Science Data," the Committee on Data of the International Council for Science (CODATA)'s "Data Science Journal" or Nature's "Scientific Data." Most of them require the data to be deposited at an approved public repository, with a digital object identifier. All of them are Open Access. Nevertheless, weather and climate data rescue activities are still often published in regular meteorological and climatological journals which, however, mostly do not yet have specific data policies. The fact that data journals have impact factors that are similar or even higher than that of climate journals suggests that the merit of climate data rescue is higher than a decade ago.

Data policies do, of course, not only concern the journal, but national policies as well as WMO recommendations. At the Seventeenth World Meteorological Congress in Geneva 2015, the WMO adopted its policy for the international exchange of climate data and products to support the implementation of the Global Framework for Climate Services (GFCS) as Resolution 60. It states that GFCS relevant data and products from specified sources (such as World Data Centres) are to be made accessible among members on a free and unrestricted basis (https://gfcs.wmo.int/sites/default/files/IBCS%20MC-3-INF-4_en.pdf).

4 | CLIMATE DATA RESCUE SERVICES

C3S DRS should address the premises above. It should support climate data rescue as a continuous effort through the Copernicus Programme, and consider climate data rescue in a distributed approach. Therefore, it needs to target individual research groups and projects, provide simple tools and aids, and at the same time guide the projects and contributions towards a common goal and a common data holding/repository. Finally, it must provide access to a community which benefits each contributing project or research group. The functioning of C3S DRS from a user perspective is shown in Figure 5. At every step from the idea of a project to data submission, C3S DRS provides support, ranging from databases, to software user support, community building, best practice reports and capacity building. The first steps are databases with existing projects

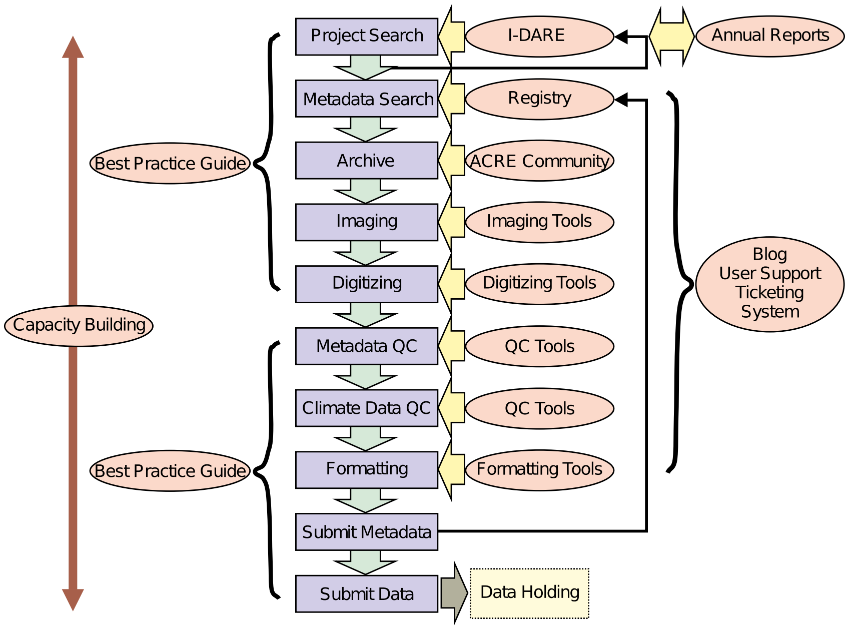

FIGURE 5 Flowchart that summarizes all phases of a data rescue project (blue rectangles). C3S DRS will provide assistance in every phase, as shown by the red circles

(I-DARE) and metadata, where crucial information can be retrieved (e.g. existing metadata on the series, links to archive holdings, contact persons and in the best cases, links to already-imaged data sheets). The annual overview of climate data rescue activities is another example, in that it should help to trigger or better embed activities.

Once digitizing is undertaken, several tools may facilitate the work surrounding the digitization process (Figure 5). Researchers will find answers to specific technical questions, such as how to digitize thermographs, barographs or pluviographs strip charts. Software tools with a free license as well as online tools will be provided that can help in every phase of digitization, from file naming of the images to the quality control of data and of the metadata (building on the software developed within the European projects ERA-CLIM and UERRA) to the organization of data rescue projects. Various format conversion tools will be provided and the data flow towards the common repository will be facilitated. Projects which are working with volunteers will find C3S DRS outreach material helpful. Moreover, an online facility for data keying will be developed. Equally important are the planned capacity building workshops, as climate data rescue projects often are needed in developing countries. Providing best practice guidelines may save time and effort, and at the same time should contribute to increased consistency across projects.

In addition to documents, the C3S DRS in conjunction with ACRE, provides a network of interested scientists. As a result, projects can then be easily embedded within an international community, with communication (blogs, newsletters), workshops, sessions at important meetings, etc. Most important in this context are links to archives, museums and libraries, interdisciplinary collaborations and end users. By being responsive to the needs of these linked groups, the service is positioned to anticipate and address future needs. In this way, climate data rescue services can make it attractive for a project to engage more widely. In turn, the project enlarges the expertise of the community and builds knowledge and "confidence" in weather and climate disciplines and in their wider value to society.

Other factors that are important to make climate data rescue attractive for scientists, but are beyond the control of C3S DRS, concern the community at large, national weather services, funding bodies and international organizations. Here, C3S DRS linkages to ACRE, and its associations with WMO, the Global Climate Observing System (GCOS), the Joint Technical Commission for Oceanography and Marine Meteorology (JCOMM), the CODATA Task Group "Data At Risk" (DAR-TG) and others, will need to come into play. Apart from the political side of data policies, the scientific side and practical aspects of data management need consideration. Climate data rescue services need to be closely linked with data holdings (e.g. Menne *et al*., 2012; Rennie *et al*., 2014; Cram *et al*., 2015; Freeman *et al*., 2017) and global repositories within C3S (Thorne *et al*., 2017) and offer these links as services to projects. In that respect, C3S DRS should channel climate data rescue projects towards common data holdings.

5 | CONCLUSIONS

Historical weather and climate data are relevant for climate sciences as well as for society in the context of climate change adaptation and mitigation, and risk assessment.

Large amounts of data and enhanced databases, vital to improving weather and climate applications and services have not yet been digitized. To promote climate data rescue, data rescue services are now being developed under C3S. In this paper, we report the results of a survey conducted among climate data rescue projects. Based on the results, we formulate three premises, from which follow requirements for such a service. We posit that climate data rescue must be seen as a continuous, long-term activity. Changes in the demand (e.g. for analysing past extreme events with respect from current to future changes in extremes) and changes in the technical possibilities (e.g. using surface pressure to obtain a complete description of the atmospheric state at high temporal and spatial scales) have led to a re-evaluation of the importance of historical data and thus the potential value of rescued data. Such changes will also occur in the future—the demand will no doubt change (e.g. a new focus on solar and wind data for planning renewable energy generation, or a stronger focus on climate impact data) and new technical possibilities will evolve (e.g. assimilation of further variables, increasing coupling of earth system components). We also postulate that a distributed approach is most suitable to the problem of rescuing climate data, and hence the task will be shared by many research projects and efforts from weather services and other institutions, through to more engagements with colleagues in the social sciences and humanities. Society is thus best served with an active climate data rescue community that provides expertise on procedures, infrastructure, tools and metadatabases. To maintain this community, ensuring scientific merit and an open data policy will make it attractive for research projects to engage in the process. Data rescue services operate as facilitators for individual projects and collectively push towards these aims. Over the last decade or so, climate data rescue efforts have led to the development of new datasets and databases. As a result, the importance of climate data rescue work has been more and more appreciated, especially as it has been increasingly used to improve the scope, quality and nature of weather and climate applications and services. These efforts must be sustained with the data rescue services under C3S DRS being a vital contribution to this effort.

ACKNOWLEDGMENTS

The authors of this paper are the team delivering Work Package 3 of the Copernicus Climate Change Service 311a Lot1 for Collection and Processing of In Situ Observations - Data Rescue. Jürg Luterbacher and Rob Allan acknowledge Climate Science for Service Partnership China (CSSP). Prof. Rob Allan acknowledges the University of Southern Queensland, Toowoomba, Australia and the Centre for Maritime Historical Studies, University of Exeter, UK where he is an Adjunct and Honorary Professor respectively. Dr. Compo was partially supported by the Physical Sciences Division of the NOAA Earth System Research Laboratory. Stefan Brönnimann acknowledges funding from the Swiss National Science Foundation (project 205121_169676).

OPEN PRACTICES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. Learn more about the Open Practices badges from the Center for Open Science: https://osf.io/tvyxz/wiki.

ORCID

Yuri Brugnara http://orcid.org/0000-0001-8427-0064

REFERENCES

Alcoforado, M.J., Vaquero, J.M., Trigo, R.M. and Taborda, J.P. (2012) Early Portuguese meteorological measurements (18th century). *Climate of the Past*, *8*, 353–371. https://doi.org/10.5194/cp-8-353-2012.

Allan, R., Brohan, P., Compo, G.P., Stone, R., Luterbacher, J. and Brönnimann, S. (2011) The international Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative. *Bulletin of the American Meteorological Society*, *92*, 1421–1425.

Allan, R., Compo, G. and Carton, J. (2011). Recovery of global surface weather observations for historical reanalyses and international users. *Eos, Transactions, American Geophysical Union*, *92*, 154. https://doi.org/10.1029/2011EO180008.

Allan, R., Endfield, G., Damodaran, V., Adamson, G., Hannaford, M., Carroll, F. *et al* (2016) Toward integrated historical climate research: the example of Atmospheric Circulation Reconstructions over the Earth. *Wiley Interdisciplinary Reviews: Climate Change*, *7*, 164–174.

Ansell, T.J., Jones, P.D., Allan, R.J., Lister, D., Parker, D.E., Brunet, M. *et al*. (2006) Daily mean sea level pressure reconstructions for the European-North Atlantic region for the period 1850–2003. *Journal of Climate*, *19*, 2717–2742.

Ashcroft, L., Gergis, J. and Karoly, D.J. (2014) A historical climate dataset for southeastern Australia, 1788–1859. *Geoscience Data Journal*, *1*, 158–178. https://doi.org/10.1002/gdj3.19.

Ashcroft, L., Allan, R.J., Bridgman, H., Gergis, J., Pudmenzky, C. and Thornton, K. (2016). Current climate data rescue activities in Australia. *Advances in Atmospheric Sciences*, *33*, 1323–1324. https://doi.org/10.1007/s00376-016-6189-5.

Ashcroft, L., Coll, J.R., Gilabert, A., Domonkos, P., Brunet, M., Aguilar, E. *et al*. (2018) A rescued dataset of sub-daily meteorological observations for Europe and the southern Mediterranean region, 1877–2012. *Earth System Science Data*, *10*(3), 1613–1635. https://doi.org/10.5194/essd-10-1613-2018.

Bonney, R., Shirk, J.L., Phillips, T.B., Wiggins, A., Ballard, H.L., Miller-Rushing, A.J. *et al*. (2014) Next steps for citizen science. *Science*, *343*, 1436–1437. https://doi.org/10.1126/science.1251554

Bosilovich, M.G., Kennedy, J.J., Dee, D., Allan, R.J. and O'Neill, A. (2013) On the reprocessing and reanalysis of observations for climate. In: Asrar, G. and Hurrell, J. (Eds.) *Climate Science for Serving Society* (pp. 51–71). Dordrecht, Germany: Springer.

Bowyer, P., Bender, S., Rechid, D. and Schaller, M. (2014). *Adapting to climate change: methods and tools for climate risk management* (p. 124). Germany: Climate Service Center.

Brönnimann, S. and Wintzer, J. (2018) Society and history imprint climate data. *Nature*, *554*, 423.

Brunet, M. and Jones, P. (2011) Data rescue initiatives: bringing historical climate data into the 21st century. *Climate Research*, *47*, 29–40. https://doi.org/10.3354/cr00960.

Brunet, M., Gilabert, A., Jones, P. and Efthymiadis, D. (2014) A historical surface climate dataset from station observations in Mediterranean North Africa and Middle East areas. *Geoscience Data Journal*, *1*, 121–128. https://doi.org/10.1002/gdj3.12.

Camuffo, D. and Bertolin, C. (2012) The earliest temperature observations in the world: the Medici Network (1654–1670). *Climatic Change*, *111*, 335–363.

Camuffo, D., Bertolin, C., Diodato, N., Cocheo, C., Barriendos, M., Dominguez-Castro, F. *et al*. (2013) Western Mediterranean precipitation over the last 300 years from instrumental observations. *Climatic Change*, *117*, 85–101.

Camuffo, D., della Valle, A., Bertolin, C., Santorelli *et al* (2017) Temperature observations in Bologna, Italy, from 1715 to 1815: a comparison with other contemporary series and an overview of three centuries of changing climate. *Climatic Change*, *142*, 7–22.

Camuffo, D. and Jones, P. (2002). *Improved understanding of past climatic variability from early daily European instrumental sources*. Dordrecht: Springer.

Challinor, A.J., Müller, C., Asseng, S., Deva, C., Nicklin, K.J., Wallach, D. *et al*. (2018) Improving the use of crop models for risk assessment and climate change adaptation. *Agricultural Systems*, *159*, 296–306. https://doi.org/10.1016/j.agsy.2017.07.010.

Compo, G.P., Whitaker, J.S. and Sardeshmukh, P.D. (2006) Feasibility of a 100-year reanalysis using only surface pressure data. *Bulletin of the American Meteorological Society*, *87*, 175–190.

Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Matsui, N., Allan, R.J., Yin, X. *et al*. (2011) The Twentieth Century Reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, *137*, 1–28.

Cram, T.A., Compo, G.P., Yin, X., Allan, R.J., McColl, C., Vose, R.S. *et al*. (2015) The International Surface Pressure Databank version 2. *Geoscience Data Journal*, *2*, 31–46. https://doi.org/10.1002/gdj3.25.

Domínguez-Castro, F., Vaquero, J.M., Rodrigo, F.S., Farrona, A.M.M., Gallego, M.C., García-Herrera, R. *et al*. (2014) Early Spanish meteorological records (1780–1850). *International Journal of Climatology*, *34*, 593–603.

Dupigny-Giroux, L.-A. and Mock, C.J. (Eds.) (2009). *Historical climate variability and impacts in North America*. Dordrecht, the Netherlands: Springer.

Freeman, E., Woodruff, S.D., Worley, S.J., Lubker, S.J., Kent, E.C., Angel, W.E. *et al*. (2017) ICOADS Release 3.0: a major update to the historical marine climate record. *International Journal of Climatology*, *37*, 2211–2232.

Fujiwara, M., Wright, J.S., Manney, G.L., Gray, L.J., Anstey, J., Birner, T. *et al*. (2017) Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems. *Atmospheric Chemistry and Physics*, *17*, 1417–1452. https://doi.org/10.5194/acp-17-1417-2017.

Griffin, R.E.M. (2006) Detection and measurement of total ozone from stellar spectra: Paper 2. Historic data from 1935–42. *Atmospheric Chemistry and Physics*, *6*, 2231–2240.

Griffin, R.E.M. (2015) When are Old Data New Data? *GeoResJ*, *6*, 92–97.

Hsu, L., Lehnert, K.A., Goodwillie, A., Delano, J.W., Gill, J.B., Tivey, M.A. *et al*. (2015) Rescue of long-tail data from the ocean bottom to the Moon: IEDA Data Rescue Mini-Awards. *GeoResJ*, *6*, 108–114. https://doi.org/10.1016/j.grj.2015.02.012.

IPCC. (2013) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J. *et al* (Eds.). Cambridge, UK and New York, NY: Cambridge University Press.

Jones, P., Davies, T., Lister, D., Slonosky, V., Jónsson, T., Bärring, L. *et al*. (1999) Monthly mean pressure reconstructions for Europe for the 1780–1995 period. *International Journal of Climatology*, *19*, 347–364. https://doi.org/10.1002/(SICI)1097-0088(19990330)19:4<347:AID-JOC363>3.0.CO;2-S.

Jourdain, S., Roucaute, E., Dandin, P., Javelle, J.-P., Donet, I., Ménassère, S. *et al*. (2015) Le sauvetage de données climatologiques anciennes à Météo-France: De la conservation des documents à la mise à disposition des données. *La Météorologie*, *8*, 47. https://doi.org/10.4267/2042/56598.

Kaspar, F., Tinz, B., Mächel, H.and Gates, L. (2015) Data rescue of national and international meteorological observations at Deutscher Wetterdienst. *Advances in Science and Research*, *12*, 57–61. https://doi.org/10.5194/asr-12-57-2015.

Klein, J., Nash, D., Priby, K., Endfield, G. and Hannaford, M. (2018) Climate, conflict and society: Changing responses to weather extremes in nineteenth century Zululand. *Environment and History*, *24*, 377–401. https://doi.org/10.3197/096734018X15137949591963.

Küttel, M., Xoplaki, E., Gallego, D., Luterbacher, J., Garcia-Herrera, R., Allan, R. *et al*. (2010) The importance of ship log data: reconstructing North Atlantic, European and Mediterranean sea level pressure fields back to 1750. *Climate Dynamics*, *34*, 1115–1128.

Kwok, R. (2017) Historical data: Hidden in the past. *Nature*, *549*, 419–421.

Laloyaux, P., de Boisseson, E., Balmaseda, M., Bidlot, J.-R., Brönnimann, S., Buizza, R. *et al*. (2018). CERA-20C: A coupled reanalysis of the Twentieth Century. *Journal of Advances in Modeling Earth System*, *10*, 1172–1195. https://doi.org/10.1029/2018MS001273.

Luterbacher, J., García-Herrera, R., Akcer-On, S., Allan, R., Alvarez-Castro, M.C., Benito, G.*et al*. (2012) A review of 2000 years of paleoclimatic evidence in the Mediterranean. In: Lionello, P. (Ed.), *The Climate of the Mediterranean region: from the past to the future* (pp. 87–185). Amsterdam, The Netherlands: Elsevier.

Meeus, S. (2018) Herbaria as Functional Trait Databases. *Biodiversity Information Science and Standards*, *2*, e25766. https://doi.org/10.3897/biss.2.25766.

Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E. and Houston, T.G. (2012) An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology*, *29*, 897–910.

Poli, P., Hersbach, H., Dee, D.P., Berrisford, P., Simmons, A.J., Vitart, F. *et al*. (2016) ERA-20C: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, *29*, 4083–4097.

Posada, R., Riede, J., Kaspar, F., Mhanda, A., Radithupa, M., Stegling, J. *et al*. (2018) Cooperation of meteorological services within SASSCAL on improving the management of observed climate data. In: Climate change and adaptive land management in southern Africa – assessments, changes, challenges, and solutions, Revermann R, Krewenka KM, Schmiedel U, Olwoch JM, Helmschrot J, Jürgens N (Eds.). *Biodiversity & Ecology*, *6*, 22–29.

Rennie, J.J., Lawrimore, J.H., Gleason, B., Thorne, P.W., Morice, C.M., Menne, M.J. *et al*. (2014) The international surface temperature initiative global land surface databank: monthly temperature data release description and methods. *Geoscience Data Journal*, *1*, 75–102. https://doi.org/10.1002/gdj3.8.

Sanchez-Romero, A., Sanchez-Lorenzo, A., Calbó, J., González, J.A. and Azorin-Molina, C. (2014) The signal of aerosol-induced changes in sunshine duration records: A review of the evidence. *Journal of Geophysical Research*, *119*, 4657–4673.

Schwander, M., Brönnimann, S., Delaygue, G., Rohrer, M., Auchmann, R. and Brugnara, Y. (2017) Reconstruction of Central European daily weather types back to 1763. *International Journal of Climatology*, *37*, 30–44.

Seidel, D.J., Gillett, N.P., Lanzante, J.R., Shine, K.P. and Thorne, P.W. (2011). Stratospheric temperature trends: Our evolving understanding. *Wiley Interdisciplinary Reviews: Climate Change*, *2*, 592–616.

Siegmund, P. (2014). *White paper on an International Data Rescue Portal (I-DARE)*, pp 16. Geneva, Switzerland: WMO.

Slonosky, V. (2014) Historical climate observations in Canada: 18th and 19th century daily temperature from the St. Lawrence Valley. *Quebec. Geoscience Data Journal*, *1*, 103–120. https://doi.org/10.1002/gdj3.11.

Stickler, A., Brönnimann, S., Valente, M.A., Bethke, J., Sterin, A., Jourdain, S. *et al*. (2014) ERA-CLIM: Historical surface and upper-air data for future reanalyses. *Bulletin of the American Meteorological Society*, *95*, 1419–1430.

Thorne, P.W., Lanzante, J.R., Peterson, T.C., Seidel, D.J. and Shine, K.P. (2011) Tropospheric temperature trends: history of an ongoing controversy. *Wiley Interdisciplinary Reviews: Climate Change*, *2*, 66–88.

Thorne, P.W., Allan, R.J., Ashcroft, L., Brohan, P., Dunn, R.J.H., Menne, M.J. *et al*. (2017) Towards an integrated set of surface meteorological observations for climate science and applications. *Bulletin of the American Meteorological Society*, *98*, 2689–2702. https://doi.org/10.1175/BAMS-D-16-0165.1.

Toride, K., Neluwala, P., Kim, H. and Yoshimura, K. (2017) Feasibility study of the reconstruction of historical weather with data assimilation. *Monthly Weather Review*, *145*, 3563–3580. https://doi.org/10.1175/MWR-D-16-0288.1.

Turner, J. and Comiso, J. (2017) Solve Antarctica's sea-ice puzzle. *Nature*, *547*, 275–277.

Vautard, R., Cattiaux, J., Yiou, P., Thépaut, J.N. and Ciais, P. (2010) Northern Hemisphere atmospheric stilling partly attributed to an increase in surface roughness. *Nature Geoscience*, *3*, 756–761.

Veale, L., Endfield, G., Davies, S., Macdonald, N., Naylor, S., Royer, M.-J. *et al*. (2017) Dealing with the deluge of historical weather data: the example of the TEMPEST database. *Geo: Geography and Environment*, *4*, e00039. https://doi.org/10.1002/geo2.39.

Wild, M. (2012). Enlightening global dimming and brightening. *Bulletin of the American Meteorological Society*, *93*, 27–37.

Wilkinson, C. and Vásquez, M. (2016) *Report on the Imaging of Historic Ice*. Vestfold Archive, Sandefjord, Norway: Meteorological and Oceanographic Data in Antarctic Waters. https://doi.org/10.13140/RG.2.2.21903.38569.

Wilkinson, C., & Vásquez, M. (2017) Report on the Imaging of Sources of Historic Ice, Meteorological and Oceanographic Data in the SouthernOcean – Åland Maritime Museum. Mariehamn, Finland. https://doi.org/10.13140/RG.2.2.24494.61763.

Williamson, F., Allan, R., Switzer, A.D., Chan, J.C.L., Wasson, R.J., D'Arrigo, R. *et al*. (2015) New directions in hydro-climatic histories: Observational data recovery, proxy records and the Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative in Southeast Asia. *Geoscience Letters*, *2*, 2. https://doi.org/10.1186/s40562-015-0018-z.

Williamson, F., Allan, R., Ren, G., Lee, T.C., Lui, W.H., Kubota, H. *et al*. (2018) Collating historic weather observations for the east Asian region: Challenges, solutions, and reanalyses. *Advances in Atmospheric Sciences*, *35*, 899–904. https://doi.org/10.1007/s00376-017-7259-z.

WMO. (2016) *Guidelines on Best Practices for Climate Data Rescue. WMO-No. 1182*, pp. 30. Geneva, Switzerland: WMO.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Brönnimann S, Brugnara Y, Allan RJ, et al. A roadmap to climate data rescue services. *Geosci Data J*. 2018;5:28–39. https://doi.org/10.1002/gdj3.56